PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G06F 3/06, 13/12

A1

(11) International Publication Number: WO 98/28679

(43) International Publication Date: 2 July 1998 (02.07.98)

(21) International Application Number: PCT/US97/22909

(22) International Filing Date: 15 December 1997 (15.12.97)

(30) Priority Data:
08/772,443 23 December 1996 (23.12.96) US

(71) Applicant: SYMBIOS, INC. [US/US]; 2001 Danfield Court, Fort Collins, CO 80525 (US).

(72) Inventors: GATES, Dennis, E.; 4893 Farmstead Court, Wichita, KS 67220 (US). GREENFIELD, Scott, E.; Apartment 1014, 1717 South Cypress, Wichita, KS 67207 (US).

(74) Agent: BAILEY, Wayne, P.; Symbios, Inc., 2001 Danfield Court, Fort Collins, CO 80525 (US).

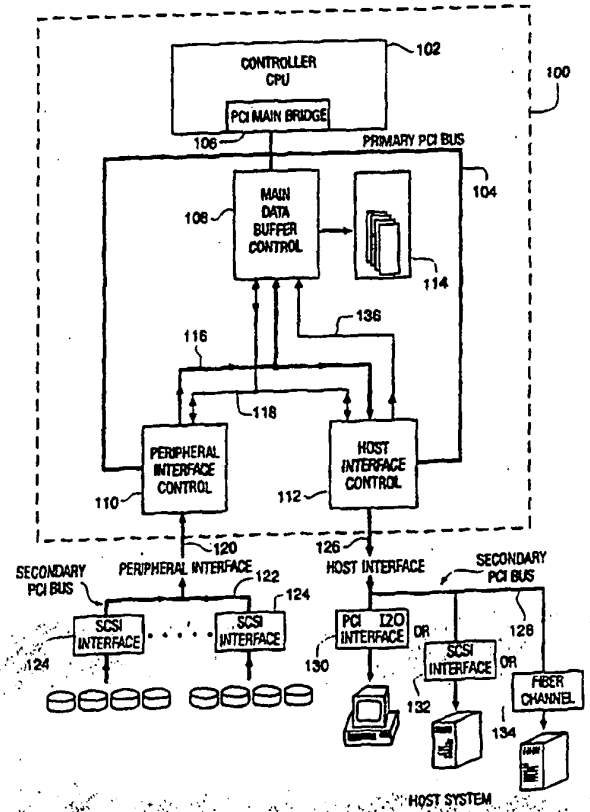
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: DATA TRANSFER WITH MIRRORING

(57) Abstract

Data transfers from the peripheral interface (120) of a disk array to a data buffer (114) are snooped to determine if the starting address of a data transfer matches an entry in a list of starting addresses for requested data. If a match is identified, third party transfer is initiated and the data is simultaneously transferred to the host interface (126) of the host system. The resulting data bandwidth is increased. A throttling/suspension mechanism can temporarily or indefinitely hold up actual data movement into the data buffer (114) to allow for temporary buffering and interface speed matching as data is transferred to the host interface (126).



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

DATA TRANSFER WITH MIRRORING

1. TECHNICAL FIELD

The present invention relates generally to transferring data in data processing systems and in particular to transferring data from a disk array to a host in a data processing system. Still more particularly, the present invention relates to transferring data from a peripheral interface for a disk array in a data processing system into a main data buffer and out to a host interface with a simple suspend/throttle control scheme.

2. DESCRIPTION OF THE RELATED ART

Redundant arrays of inexpensive disks (RAID) such as small SCSI hard disks, have been found to be suitable alternatives to large capacity, single magnetic media. Such arrays appear to the host system as a single media, but provide a very high data transfer rate through a technique called "striping." The arrays also provide improved reliability, scalability, data availability and power consumption over large magnetic disks.

Control of the array for READ, WRITE, and other operations may be performed by the host, but is typically effected by a controller. Controllers generally utilize a data buffer between the host and the array, transferring data from the array to the data buffer and from the data buffer to the host.

Current implementations of buffered controllers support non-cached disk READ operations by performing high-speed data movement between a peripheral interface and a main data buffer and between a main data buffer and a host interface. The total data transfer is generally accomplished in two distinct operations: the inbound transfer places data read from the peripheral interface of the array into the main data buffer, while the outbound transfer removes data from the main data buffer and forwards it to the host interface.

It would be advantageous to transfer data from the peripheral interface of the array into the data buffer and out to the host interface in a single operation, thereby achieving a higher effective data bandwidth as well as improving the host read data response time.

3. SUMMARY OF THE INVENTION

Data transfers from the peripheral interface of a disk array to a data buffer are snooped to determine if the starting address of a data transfer matches an entry in a list of starting addresses for requested data. If a match is identified, third party transfer is initiated and the data is simultaneously transferred to the host interface of the host system. The resulting data bandwidth is increased. A throttling/suspension mechanism can temporarily or indefinitely hold up actual data movement into the data buffer to allow for temporary buffering and interface speed matching as data is transferred to the host interface.

4. BRIEF DESCRIPTION OF THE DRAWINGS

The novel features believed characteristic of the invention are set forth in the appended claims. The invention itself, however, as well as a preferred mode of use, further objectives and advantages thereof, will best be understood by reference to the following detailed description of an illustrative embodiment when read in conjunction with the accompanying drawings, wherein:

Figure 1 depicts a data processing system employing a disk array and a controller in accordance with a preferred embodiment of the present invention;

Figure 2 is a block diagram illustrating further details of the main data buffer control within the controller depicted in **Figure 1**;

Figure 3 depicts further details of the peripheral interface control within the controller depicted in **Figure 1**;

Figure 4 is a block diagram illustrating further details of the host interface control within the controller depicted in **Figure 1**;

Figure 5 depicts a high level flowchart for a process in the snooping mechanism for determining if a requested data transfer is occurring in accordance with a preferred embodiment of the present invention;

Figure 6 is a high level flowchart for a process of throttling/suspending data transfer into the controller main data buffer in accordance with a preferred embodiment of the present invention; and

Figures 7A-7C depict timing diagrams demonstrating the effect of a throttling/suspension mechanism on data transfer into the data buffer.

5. DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

With reference now to the figures, and in particular with reference to **Figure 1**, a block diagram of a data processing system employing a disk array and a controller in accordance with a preferred embodiment of the present invention is depicted. Controller **100** comprises controller CPU **102** linked to primary PCI bus **104** by PCI main bridge **106**. Primary PCI bus **104** also connects main data buffer control **108** to both peripheral interface control **110** and host interface control **112**. Main data buffer control **108** is connected to, and controls data transfer to and from, data buffer **114**. Main data buffer control **108** is also

linked to both peripheral interface control 110 and host interface control 112 by buffer data bus 116 and main control bus 118.

In operation, peripheral interface control 110 controls the transfer of data from peripheral interface 120 to secondary PCI bus 122, which is connected, for example, by a plurality of SCSI interfaces 124 to the storage media of the disk array. Data read from the storage media is transferred by peripheral interface control 110 into data buffer 114 via buffer data bus 116 and main data buffer control 108. The data may be transferred out of data buffer 114 to host interface control 112 via main data buffer control 108 and buffer data bus 116. Buffer data bus 116 is used to transfer the data itself while main control bus 118 is used to control the transfer and for signaling data request/acknowledge.

Host interface control 112 controls the transfer of data through host interface 126 to secondary PCI bus 128. Secondary PCI bus 128 is connected to the host system by an appropriate interface such as PCI interface 130, SCSI interface 132, or fiber channel interface 134.

As described below, host interface control 112 contains a snooping mechanism to detect transfers of requested data from peripheral interface control 110 to main data buffer control 108. Host interface control 112 also contains a throttling/suspension mechanism, which employs throttling control bus 136 between host interface control 112 and main data buffer control 108 to act as a pacing agent for actual data transfer.

Although depicted separately for clarity in explaining the present invention, those skilled in the art will recognize that buffer data bus 116, main control bus 118,

and throttle control bus **136** may employ a single set of conductors and/or pins and that main data buffer control **108**, data buffer **114**, peripheral interface control **110**, and host interface control **112** may be implemented as a single integrated circuit.

While the exemplary embodiment described herein employs a PCI bus, those skilled in the art will recognize that the principles disclosed are applicable to any data transfer from a source to a destination which employs intermediate buffering, regardless of the bus architecture selected.

Referring to **Figure 2**, further details of the main data buffer control within the controller depicted in **Figure 1** are illustrated in a block diagram. Main data buffer control **108** regulates data movement from buffer data bus **116** to the actual data buffer. Buffer timing and control logic unit **200**, connected to main control bus **118** and throttling control bus **136**, contains sequential state machine logic controlling both the timing and assertion of signals to the specific buffer devices and the transfer of data from buffer data bus interface unit **202** to data checking and data path unit **204**. Buffer timing and control logic unit **200** handles buffer refresh requirements, timing requirements for various cycle control signals, and address/data transfer requirements relative to the buffer devices in the main data buffer. These tasks are common to most memory control logic such as that used for DRAMs and are not unique to the present invention.

Data checking and data path unit **204** contains registers and steering logic necessary to support movement of data from buffer data bus interface **202** to the data buffer. Control signals for these registers and steering logic are received from buffer timing and control logic unit **200** via signal lines **206**. Data is forwarded to the

data buffer from data checking and data path unit **204** via data lines **212**.

Buffer data bus interface **202** contains hardware providing an immediate interface of control and data bits from buffer data bus **116**. This hardware typically includes at least one level of buffering or temporary storage of data bits for transfer from buffer data bus **116** to data checking and data path unit **204**. Controls signals for this hardware are received from buffer timing and control logic unit **200** via signal lines **208**, while the actual data transfer to data checking and data path unit **204** flows via data lines **210**.

With the exception of the input for throttle control bus **136** and attendant circuitry for responding appropriately to signals on throttle control bus **136**, main data buffer control **108** may be constructed in accordance with known embodiments of disk array controllers.

With reference now to **Figure 3**, further details of the peripheral interface control within the controller depicted in **Figure 1** are depicted. Peripheral interface control **110** controls transfer of data from peripheral interface **120** to buffer data bus **116**. Peripheral bus interface unit **300** contains hardware providing an immediate interface of control and data bits from peripheral interface **120**. This hardware typically includes at least one level of buffering of data bits for transfer to peripheral FIFO buffer and controls **302**. Peripheral bus interface unit **300** receives control signals from peripheral interface control unit **304** via signal lines **306** and forwards data to peripheral FIFO buffer and controls **302** via data lines **308**.

Peripheral interface control unit **304** contains sequential state machine logic controlling data movement between peripheral interface **120** into peripheral bus

interface unit 300 as well as data transfer from peripheral bus interface unit 300 to peripheral FIFO buffer and controls 302.

Peripheral FIFO buffer and controls 302 contains a FIFO (first-in, first-out) buffer and data path steering logic necessary to support efficient movement of input data from peripheral bus interface unit 300 to buffer bus interface unit 310. Control signals for this buffer and steering logic are received from peripheral bus interface unit 300 via signal line 312 and from buffer interface control unit 314 via signal line 316. Data from peripheral FIFO buffer and controls 302 is forwarded to buffer bus interface unit 310 via data lines 318.

Buffer interface control unit 314 contains sequential state machine logic controlling movement of data from peripheral FIFO buffer and controls 302 into buffer bus interface unit 310 and from buffer bus interface unit 310 to buffer data bus 116. Buffer interface control unit 314 sends control signals to buffer bus interface unit 310 via signal lines 320. Buffer bus interface unit 310 provides an immediate interface for control and data bits being forwarded to buffer data bus 116 and typically includes a minimum of one level of buffering of data bits for transfer from peripheral FIFO buffer and controls 302 to buffer data bus 116.

The components of peripheral interface control 110 may be constructed in accordance with known embodiments of disk array controllers. The tasks described are common to data transfers from disk arrays and are not unique to the present invention.

Referring to Figure 4, further details of the host interface control within the controller depicted in Figure 1 are illustrated in a block diagram. Host interface

control 112 controls data transfer from buffer data bus 116 to host interface 126. Buffer bus interface unit 400 contains hardware providing an immediate interface controlling data movement from buffer data bus 116 to host FIFO buffer and controls 402. This hardware typically includes a minimum of one level of buffering of data bits transferred from buffer data bus 116 to host FIFO buffer and controls 402. Control signals for this hardware are received from buffer interface control unit 404 via signal lines 406, while data is transferred from buffer data bus 116 to host FIFO buffer and controls 402 via data line 408.

Host FIFO buffer and controls 402 contains a FIFO buffer and data path steering logic necessary to support efficient movement of input data from buffer bus interface unit 400 to host bus interface unit 410. Controls signals for the buffer and steering logic are received from buffer interface control unit 404 via signal line 412 and from host interface control unit 414 via signal line 416. Data is transferred from host FIFO buffer and controls 402 to host bus interface unit 410 via data lines 418.

Host interface control unit 414 contains sequential state machine logic controlling movement of data from host FIFO buffer and controls 402 to host bus interface unit 410 and from host bus interface unit 410 to host interface 126. Control signals are sent from host interface control unit 414 to host bus interface unit 410 via signal lines 420. Host bus interface unit 410 contains hardware providing an immediate interface of control and data bits to host interface 126. This hardware typically includes at least one level of buffering of data bits transferred from host FIFO buffer and controls 402 to host interface 126.

With the exception of snooping circuit and queue control logic 422, throttle and suspension logic 424, and

modifications to buffer interface control unit 404, as described below, the components of host interface control 112 may be constructed in accordance with known embodiments of disk array controllers. The tasks described are common to data transfers from disk arrays and are not unique to the present invention.

Buffer interface control unit 404 contains sequential state machine logic which, in conjunction with snooping circuit and queue control logic 422 and throttle and suspension logic 424, controls movement of data from buffer data bus 116 into buffer bus interface unit 400 as well as movement of data from buffer bus interface unit 400 into host FIFO buffer and controls 402. Buffer interface control unit 404 also manages throttle control bus 136.

Snooping circuit and queue control logic 422 contains hardware enabling host interface control 112 to participate in data transfers occurring on buffer data bus 116. Snooping circuit and queue control logic 422 includes a hardware queueing mechanism containing a list of potential starting addresses for data transfers on buffer data bus 116. This list of starting addresses can be polled sequentially to allow an effective comparison of all actual data transfer starting addresses to all starting address entries in the queue.

If snooping circuit and queue control logic 422 detects a match between an actual data transfer starting address and a starting address in the queue, it signals buffer interface control unit 404 to initiate transfer of the data on buffer data bus 116 into buffer bus interface unit 400. Thus requested data may be obtained for host interface 126 directly from the peripheral interface control via buffer data bus 116, without first storing the data in the data buffer and subsequently transferring the

data via host interface control **112** to host interface **126**.

Snooping circuit and queue control logic **422** also includes a device enabling throttle and suspension logic **424** to initiate monitoring and passive control of the current handshake of data transferred from the peripheral interface control to buffer data bus **116** and received by the main data buffer control. Snooping circuit and queue control logic **422** also contains a mechanism determining when a continuation or suspension of a previously queued data transfer is required. This requires some data transfer length detection or counter logic as well as holding registers for suspended or pending transfers considered to be in progress but not currently transferring on buffer data bus **116**.

Snooping circuit and queue control logic **422** provides an input to buffer interface control unit **404** via signal line **426** and receives an input from buffer interface control unit **404** via signal line **428**. Snooping circuit and queue control logic **422** receives addresses from the CPU on primary PCI bus **104**.

Throttle and suspension logic **424** contains logic for determining whether a current data transfer on buffer data bus **116** should be throttled or suspended. Throttle and suspension logic **424** provides an input to buffer interface control unit **404** via signal line **430** for use in managing throttle control bus **136**. Throttle and suspension logic **424** receives an input from buffer interface control unit **404** via signal line **432** used in determining if throttling or suspension is required.

With reference now to **Figure 5**, a high level flowchart is depicted for a process in the snooping mechanism for determining if a requested data transfer is occurring in accordance with a preferred embodiment of the present

invention. The process is preferably performed by snooping control circuitry included in the host interface control of a disk array controller. The snooping control circuity includes a queue listing possible data transfer starting addresses, and in particular the starting addresses of requested data transfers.

The process begins at step 500, which depicts the beginning of a data transfer into the main data buffer via the buffer data bus. The process passes next to step 502, which illustrates a comparison of the starting address of the data transfer on the buffer data bus to the list of starting addresses in the queue, and then to step 504, which depicts a determination of whether the starting address of the data transfer matched a starting address in the queue. If not, the process proceeds to step 506, which illustrates the process becoming idle until the beginning of the next data transfer on the buffer data bus. If so, however, the process proceeds instead to step 508, which depicts initiation of third party or "fly-by" data transfer and begin receiving the data as it is transferred from the peripheral interface control to the main data buffer control and data buffer. The process then passes to step 510, which depicts monitoring the data transfer to determine is throttling or suspension is required to allow temporary buffering and interface speed matching as data is transferred to the host interface.

Referring to **Figure** 6, a high level flowchart for a process of throttling/suspending data transfer into the controller main data buffer in accordance with a preferred embodiment of the present invention is illustrated. The throttling/suspension is preferably effected by a signal from the host interface control to the main data buffer control on the throttle control bus is response to a determination by the throttling mechanism in the host interface control that throttling/suspension is required.

The process begins at step 600, which depicts the

beginning of a third party data transfer into the host interface control of data being transferred from the peripheral interface control to the main data buffer. The process then passes to step 602, which illustrates a determination of whether throttling or suspension is necessary due to the host interface buffer filling to a level and/or at a rate which indicates the host system requires additional clock cycles to receive the data already transferred from the peripheral interface control to the main data buffer control. If the host interface buffer is filling, the process proceeds to step 604, which depicts signaling a throttle of the data transfer on the throttle control bus, causing the main data buffer control to assert a signal disabling the data request. The process next passes to step 606, which illustrates a determination of whether the host interface buffer is clearing in response to transfer of data to the host system. If not, the process loops back to step 606 continually until the host interface buffer is sufficiently cleared. If so, however, the process proceeds instead to step 608, described below.

Referring again to step 602, if throttling or suspension is not required, the process passes to step 608, which depicts a determination of whether the data transfer is complete. If not, the process loops back to step 602 to again determine if throttling or suspension is required. if so, however, the process proceeds to step 610, which illustrates the process becoming idle until the next third party data transfer to the host interface buffer begins.

With reference now to **Figures 7A-7C**, timing diagrams demonstrating the effect of a throttling/suspension mechanism on data transfer into the data buffer are depicted. The "valid" notation shown for the ADDRESS signal in each timing diagram indicates that a starting address matching an entry in the queue has been detected.

Implementation of the START and END signals is a matter of design choice since these signals merely indicate a bounding condition for a given data transfer. The actual points of data transfer are indicated by the numbers in circles. In the depicted example, data is transferred only on the rising clock edge if DATA ACKNOWLEDGE is ON (high). DATA ACKNOWLEDGE can be turned ON for the next rising clock edge if, for the current rising clock edge, DATA REQUEST is ON (high) and DISABLE DATA REQUEST is OFF (low).

Figure 7A depicts a timing diagram for an unthrottled, unsuspended data transfer. **Figure 7B** depicts a timing diagram for a data transfer which is throttled for two clock cycles and then resumed. **Figure 7C** depicts a timing diagram for a data transfer which was suspended (or indefinitely throttled) before any data was actually transferred.

Controlling data transfers between a host system and a disk array in accordance with the present invention results in efficient data transfer and improved data bandwidth. The present invention also improves the host read data response time.

The description of the preferred embodiment of the present invention has been presented for purposes of illustration and description, but is not intended to be exhaustive or limit the invention in the form disclosed. Many modifications and variations will be apparent to those of ordinary skill in the art. The embodiment was chosen and described in order to best explain the principles of the invention and the practical application to enable others of ordinary skill in the art to understand the invention for various embodiments with various modifications as are suited to the particular use contemplated.

CLAIMS:

What is claimed is:

1. A disk array controller, comprising:

a peripheral interface control regulating data transfers from a peripheral to a bus;

a data buffer control regulating data transfers from the bus to a data buffer;

a host interface control regulating data transfers from the bus to a host system;

a snooping circuit monitoring data transfers to the bus to determine if a current data transfer on the bus matches a requested data transfer,

wherein, in response to a determination that the current data transfer on the bus matches a requested data transfer, the host interface control initiates a transfer of data from the bus to the host system.

2. The disk array controller of claim 1, further comprising:

a throttle circuit comparing a data transfer on the bus to a data transfer from the bus to the host system, the throttle circuit signaling a suspension of the data transfer on the bus in response to determining that the data transfer from the bus to the host system is slower than the data transfer on the bus.

3. The disk array controller of claim 1, wherein the peripheral interface control further comprises:

a buffer;

a first interface unit transferring data between the peripheral and the buffer;

a first control unit controlling the first interface unit;

a second interface unit transferring data from the buffer to the bus; and

a second control unit controlling the second interface

unit.

4. The disk array controller of claim 1, wherein the data buffer control further comprises:

a first interface unit transferring data between the bus and the data buffer;

a first control unit controlling the first interface unit; and

a steering unit directing data transfers from the first interface unit to the data buffer.

5. The disk array controller of claim 1, wherein the host interface control further comprises:

a buffer;

a first interface unit transferring data between the bus and the buffer;

a first control unit controlling the first interface unit;

a second interface unit transferring data from the buffer to the host system;

a second control unit controlling the second interface unit, wherein the second control unit transmits a throttle signal to the data buffer control; and

the snooping circuit, wherein the snooping circuit receives an input signal from the second control unit and transmits an output signal to the second control unit,

the snooping circuit including a memory containing a list of requested data transfers,

the snooping circuit comparing a current data transfer on the bus to the list of requested data transfers,

the snooping circuit, in response to determining that the current data transfer matches a requested data transfer in the list, transmitting a first signal to the first control unit, and

the snooping circuit, in response to determining that the current data transfer does not match a requested data transfer in the list, transmitting a second signal to the

first control unit.

6. The disk array controller of claim 5, wherein the first control unit, in response to receiving the first signal from the snooping circuit, initiates a data transfer by the first interface unit from the bus to the buffer.

7. The disk array controller of claim 5, wherein the host interface control further comprises:

a throttle circuit receiving an input signal from the first control unit and transmitting an output signal to the first control unit,

the throttle circuit comparing a first data transfer rate for a data transfer by the first interface unit to a second data transfer rate for a data transfer by the second interface unit,

the throttle circuit, in response to determining that the first data transfer rate does not exceed the second data transfer rate, transmitting a first signal to the first control unit, and

the throttle circuit, in response to determining that the first data transfer rate exceeds the second data transfer rate, transmitting a second signal to the first control unit.

8. The disk array controller of claim 7, wherein the first control unit, in response to receiving the second signal from the throttle circuit, transmits a throttle signal to the data buffer control.

9. A bus controller, comprising:

first control means for regulating data transfers from a peripheral to a bus;

second control means for regulating data transfers from the bus to a data buffer;

third control means for regulating data transfers from the bus to a host system;

comparison means for determining if a data transfer on the bus matches a requested data transfer,

wherein, in response to a determination that the current data transfer on the bus matches a requested data transfer, the bus controller initiates a transfer of data from the bus to the host system.

10. The bus controller of claim 9, wherein the comparison means forms a first comparison means, the bus controller further comprising:

second comparison means for comparing a data transfer on the bus to a data transfer from the bus to the host system; and

fourth control means for signaling a suspension of the data transfer on the bus in response to determining that the data transfer from the bus to the host system is slower than the data transfer on the bus.

11. The bus controller of claim 9, wherein the first control means further comprises:

buffer means for buffering data transfers from the peripheral to the bus;

first interface means for transferring data from the peripheral to the buffer means;

first interface control means for controlling the first interface means;

second interface means for transferring data from the buffer means to the bus; and

second interface control means for controlling the second interface means.

12. The bus controller of claim 9, wherein the second control means further comprises:

first interface means for transferring data between the bus and the data buffer;

first interface control means for controlling the first interface means; and

data steering means for directing data transfers from the first interface means to the data buffer.

13. The bus controller of claim 8, wherein the third control means further comprises:

buffer means for buffering data transfers from the bus to the host system;

first interface means for transferring data between the bus and the buffer means;

first interface control means for controlling the first interface means;

second interface means for transferring data from the buffer means to the host system;

second interface control means for controlling the second interface means, wherein the second interface control means transmits a throttle signal to the second control means;

memory means for holding a list of requested data transfers; and

the comparison means, wherein the comparison means receives an input signal from the second interface control means and transmits an output signal to the second interface control means,

the comparison means comparing a current data transfer on the bus to the list of requested data transfers,

the comparison means, responsive to determining that the current data transfer matches a requested data transfer in the list, transmitting a first signal to the first interface control means, and

the comparison means, responsive to determining that the current data transfer does not match a requested data transfer in the list, transmitting a second signal to the first interface control means.

14. The bus controller of claim 13, wherein the first interface control means, responsive to receiving the first signal from the comparison means, initiates a data transfer

by the first interface means from the bus to the buffer means.

15. The bus controller of claim 13, wherein the comparison means forms a first comparison means and the host interface control further comprises:

second comparison means for comparing a first data transfer rate for a data transfer by the first interface means to a second data transfer rate for a data transfer by the second interface means,

the second comparison means receiving an input signal from the first interface control means and transmitting an output signal to the first interface control means,

the second comparison means, in responsive to determining that the first data transfer rate does not exceed the second data transfer rate, transmitting a first signal to the first interface control means, and

the second comparison means, responsive to determining that the first data transfer rate exceeds the second data transfer rate, transmitting a second signal to the first interface control means.

16. The bus controller of claim 15, wherein the first interface control means, responsive to receiving the second signal from the second comparison means, transmits a throttle signal to the second control means.

17. A controller, comprising:

a peripheral interface;

a data buffer;

a host system interface; and

a bus transferring data and commands between the peripheral interface, the data buffer and the host system interface,

wherein data transfers on the bus between the peripheral interface and the data buffer are snooped, and, in response to determining that a current data transfer

between from the peripheral interface to the data buffer matches a requested data transfer, the current data transfer is simultaneously received by the data buffer and the host system interface.

18. The controller of claim 17, wherein the data transfers on the bus are snooped by a host interface control.

19. The controller of claim 17, wherein the controller has a first mode of operation in which data transfers between the peripheral interface and the data buffer progress and a second mode of operation wherein data transfers between the peripheral interface and the data buffer are suspended.

20. A data processing system, comprising:

a disk array;

a host system requesting and receiving data transfers originating from the disk array;

a bus transferring data and commands between the disk array and the host system, wherein data transfers on the bus are snooped, and, in response to determining that a current data transfer matches a requested data transfer, the current data transfer is received by the host system without buffering.

21. The data processing system of claim 20, further comprising:

a peripheral interface control connecting the disk array to the bus;

a data buffer control connecting the bus to a data buffer;

a host interface control connecting the bus to a host system;

a snooping circuit monitoring data transfers on the bus to determine if the current data transfer matches a requested data transfer.

22. The data processing system of claim 21, wherein the snooping circuit forms a portion of the host interface control.

23. The data processing system of claim 21, wherein the snooping circuit includes a memory containing a list of requested data transfers.

24. The data processing system of claim 20, further comprising:

a throttle circuit comparing a data transfer on the bus to a data transfer from the bus to the host system, the throttle circuit signaling a suspension of the data transfer on the bus in response to determining that the data transfer from the bus to the host system is slower than the data transfer on the bus.

25. The data processing system of claim 24, wherein the throttle circuit forms a portion of the host interface control.

26. A method of transmitting data from a disk array to a host system, comprising:

transmitting a data packet from the disk array onto a bus;

transmitting the data packet to a data buffer from the bus;

monitoring an attribute of the data packet to determine if the data packet is a requested data packet; and

responsive to determining that the data packet is a requested data packet, transmitting the data packet to the host system from the bus simultaneously with the transmission of the data packet to the data buffer.

27. The method of claim 26, wherein the step of monitoring an attribute of the data packet further comprises

maintaining a list containing requested data transfers and corresponding attributes and comparing the attribute of the data packet to attributes contained in the list.

28. The method of claim 26, wherein the step of monitoring an attribute of the data packet further comprises determining the source address of the data packet.

29. The method of claim 26, wherein

the step of transmitting a data packet from the disk array onto a bus further comprises transmitting a plurality of related data packets from the disk array onto the bus, and

the step of transmitting the data packet to the host system from the bus further comprises transmitting the plurality of data packets from the bus to the host system.

30. The method of claim 29, further comprising:

comparing a first transfer rate for transmission of the plurality of data packets from the disk array onto the bus to a second transfer rate for transmission of the plurality of data packets from the bus to the host system; and

responsive to determining that the second transfer rate is slower than the first transfer rate, suspending transmission of the plurality of data packets from the disk array onto the bus.

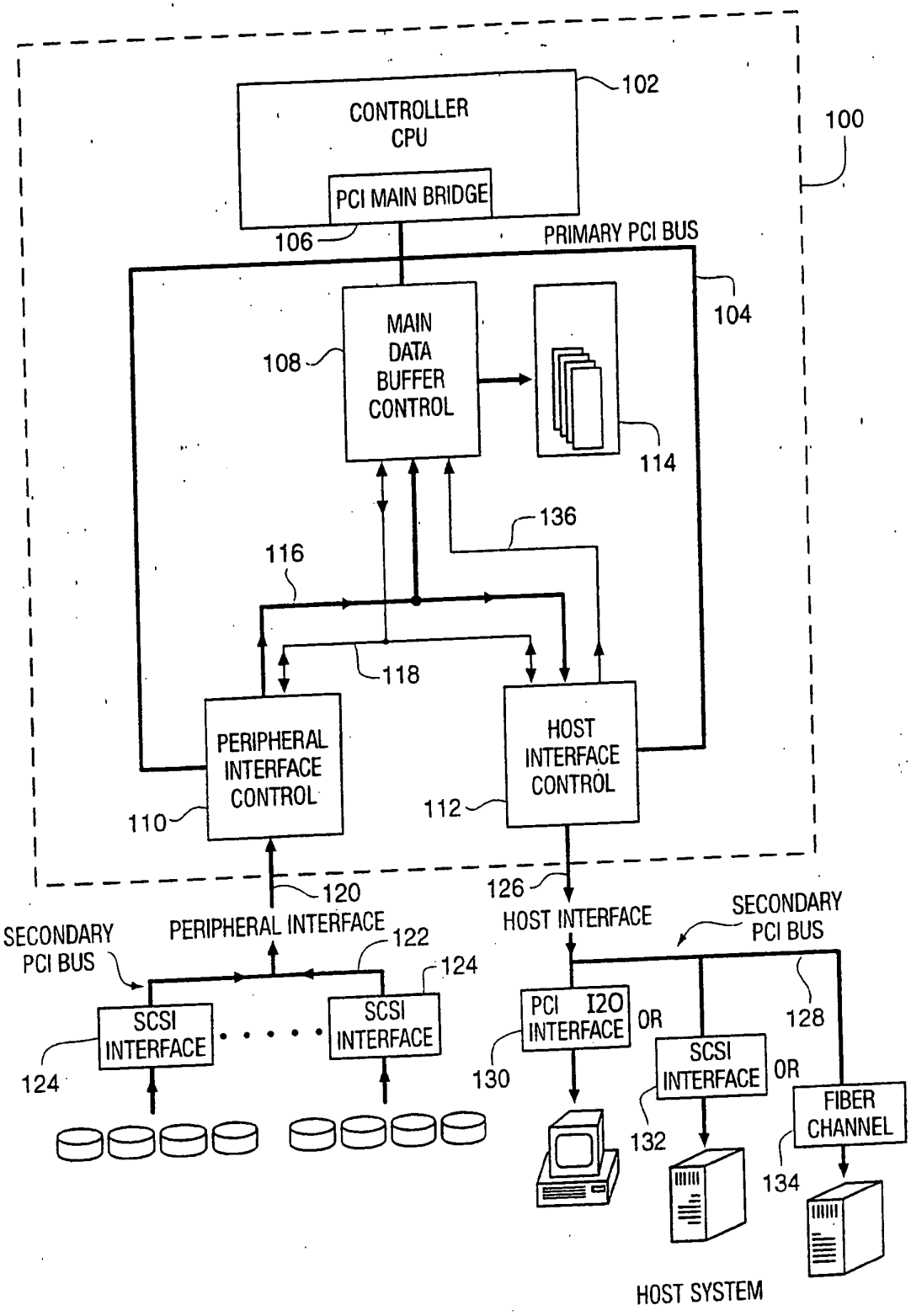

CONTROLLER CPU
102
100
PCI MAIN BRIDGE
106
PRIMARY PCI BUS
104
MAIN DATA BUFFER CONTROL
108
114
136
116
118
PERIPHERAL INTERFACE CONTROL
110
HOST INTERFACE CONTROL
112
120
126
SECONDARY PCI BUS
PERIPHERAL INTERFACE
122
124
HOST INTERFACE
SECONDARY PCI BUS
SCSI INTERFACE
SCSI INTERFACE
124
PCI I2O INTERFACE
OR
128
130
SCSI INTERFACE
OR
FIBER CHANNEL
132
134
HOST SYSTEM

FIG. 1

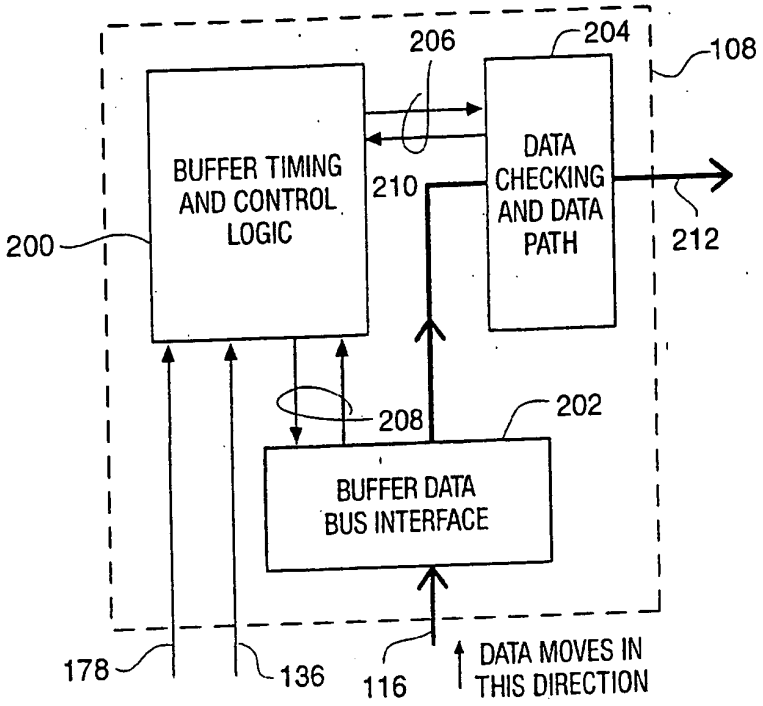

206
204
108
BUFFER TIMING AND CONTROL LOGIC
DATA CHECKING AND DATA PATH
210
200
212
208
202
BUFFER DATA BUS INTERFACE
178
136
116
DATA MOVES IN THIS DIRECTION

FIG. 2

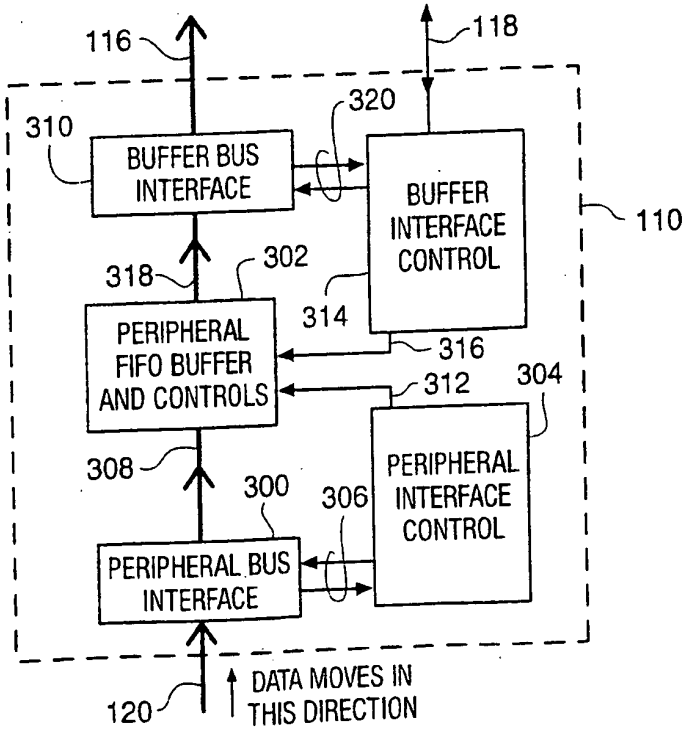

116
118
320
310
BUFFER BUS INTERFACE
BUFFER INTERFACE CONTROL
110
318
302
314
PERIPHERAL FIFO BUFFER AND CONTROLS
316
312
304
308
300
306
PERIPHERAL INTERFACE CONTROL
PERIPHERAL BUS INTERFACE
120
DATA MOVES IN THIS DIRECTION

FIG. 3

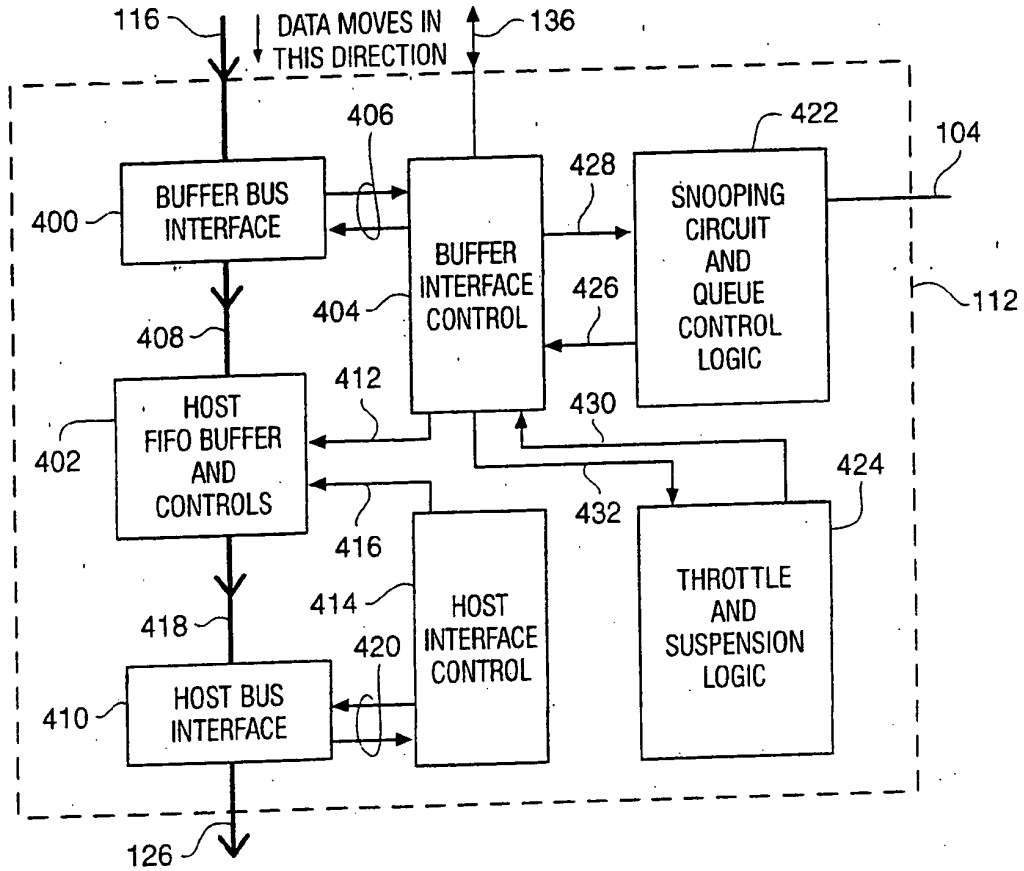

116
DATA MOVES IN THIS DIRECTION
136
406
422
104
428
400
BUFFER BUS INTERFACE
BUFFER INTERFACE CONTROL
SNOOPING CIRCUIT AND QUEUE CONTROL LOGIC
426
404
112
408
412
430
402
HOST FIFO BUFFER AND CONTROLS
424
432
416
414
HOST INTERFACE CONTROL
THROTTLE AND SUSPENSION LOGIC
418
420
410
HOST BUS INTERFACE
126

FIG. 4

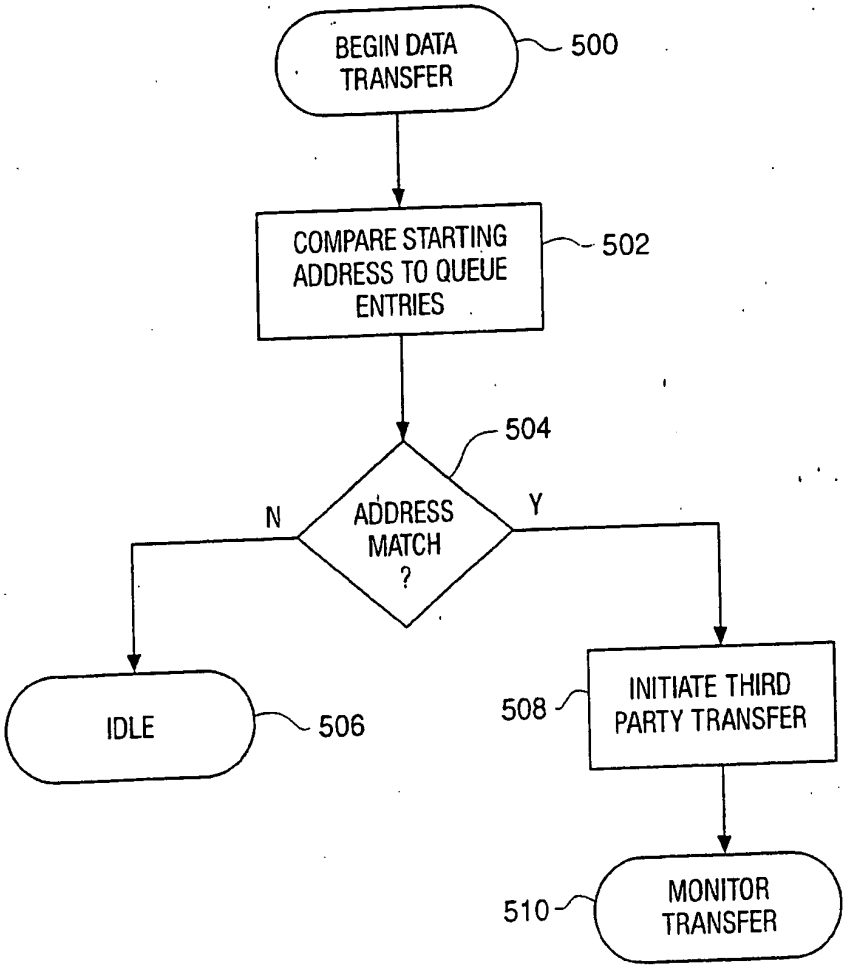

BEGIN DATA TRANSFER
500
COMPARE STARTING ADDRESS TO QUEUE ENTRIES
502
504
ADDRESS MATCH ?
N
Y
IDLE
506
508
INITIATE THIRD PARTY TRANSFER
MONITOR TRANSFER
510

FIG. 5

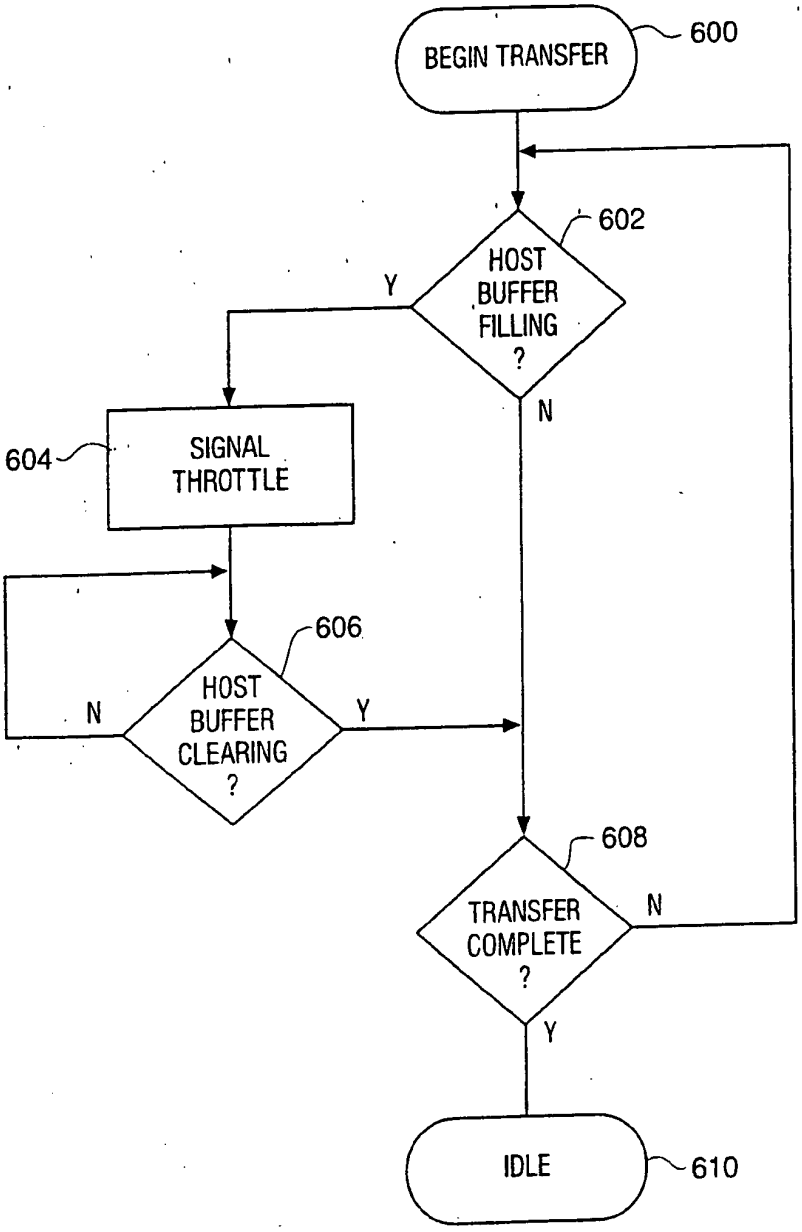

BEGIN TRANSFER
600
602
HOST BUFFER FILLING ?
Y
N
604
SIGNAL THROTTLE
606
HOST BUFFER CLEARING ?
N
Y
608
TRANSFER COMPLETE ?
N
Y
IDLE
610

FIG. 6

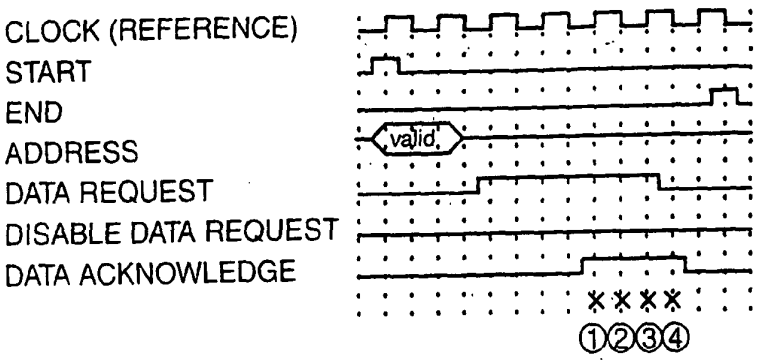


FIG. 7A

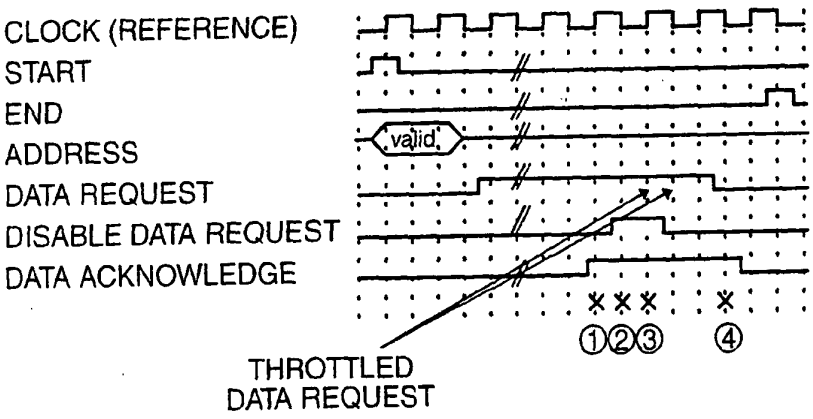


FIG. 7B

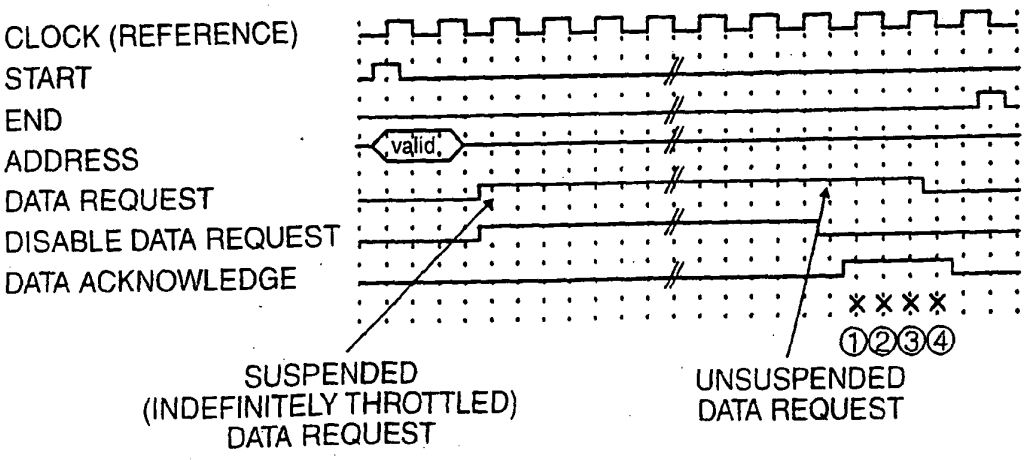


FIG. 7C

INTERNATIONAL SEARCH REPORT

Inte. .onal Application No
PCT/US 97/22909

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G06F3/06 G06F13/12

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G06F

Documentation searched other than minimumdocumentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 530 830 A (IWASAKI HIDEHIKO ET AL) 25 June 1996 see column 2, line 53 - column 3, line 43; figure 1	1,3-5,9, 11,12, 17,20,26
A	US 5 353 415 A (WOLFORD JEFF W ET AL) 4 October 1994 see abstract see column 7, line 16 - line 32; figure 3	1,9,17, 20,26
A	US 5 548 788 A (MCGILLIS JAMES M ET AL) 20 August 1996 see column 1, line 48 - column 3, line 40; figure 1	1,9,17, 20,26

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publicationdate of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of theinternational search	Date of mailing of the international search report
23 April 1998	04/05/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Moens, R

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte. .onal Application No
PCT/US 97/22909

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5530830 A	25-06-96	JP 6175786 A	24-06-94
US 5353415 A	04-10-94	AU 5403394 A	26-04-94
		WO 9408297 A	14-04-94
US 5548788 A	20-08-96	NONE	